

머신러닝 & 딥러닝 OT

AI 학술동아리 <MLP>

- Index

0. 참고(책)

1. AI(머신러닝 & 딥러닝)란?

2. 환경 세팅

3. 머신러닝 기초

k-NN 알고리즘

0. 참고

혼자 공부하는 머신러닝 + 딥러닝

인공지능 공부가 정말 처음일 때
어려운 수식에 지쳤을 때
쉬운 그림과 실전 예제로 공부하고 싶을 때

박해선 지음

★★★★★

인공지능

혼자 공부하는 일대 능숙한 사람 혹은 그런 무리를 일컫는 신조어

(유미어) 혼공족, 혼공파, 혼공자, 혼공비

구글 코랩으로
환경 설정 없이
실행 가능

유튜브 강의
발책
용어 노트

TALK
저자와 함께하는
오픈채팅

<http://bit.ly/tensor-chat>

한빛미디어

1. AI란? (AI, ML, DL)



1-1. ML(머신러닝)이란? (ex : Titanic)

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

생존여부가 주어진 train data

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S

생존여부가 주어지지 않은 test data

train data 학습
→ **test data 예측**

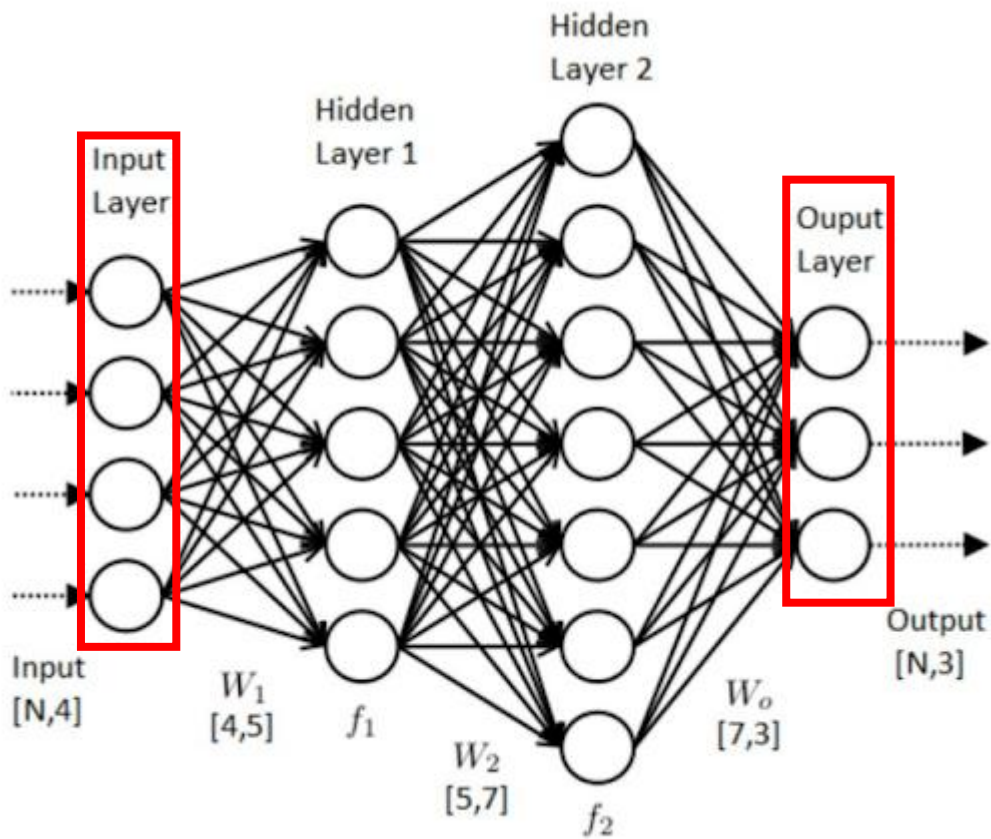
1-1. ML이란?

- 머신러닝 라이브러리 : `sckit-learn`(사이킷 런)



1-2. DL(딥러닝)이란?

train data 학습
→ **test data 예측**은 동일



ANN(Artificial Neural Network)(인공신경망)

1-2. DL이란?

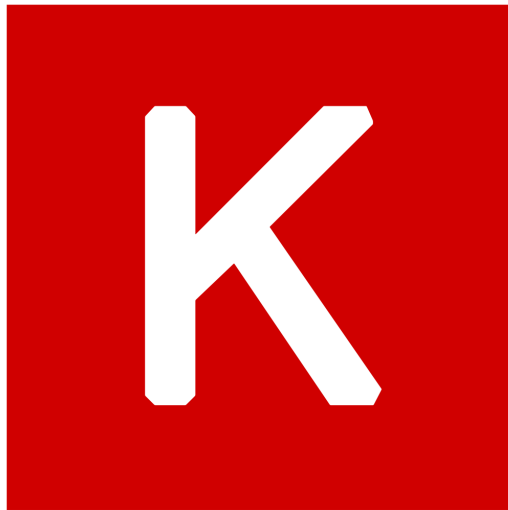
- 딥러닝 라이브러리 : TensorFlow(Keras), PyTorch



TensorFlow



PyTorch



현재(2023년) 기준 PyTorch가 가장 인기 있음

ML, DL 간단 요약

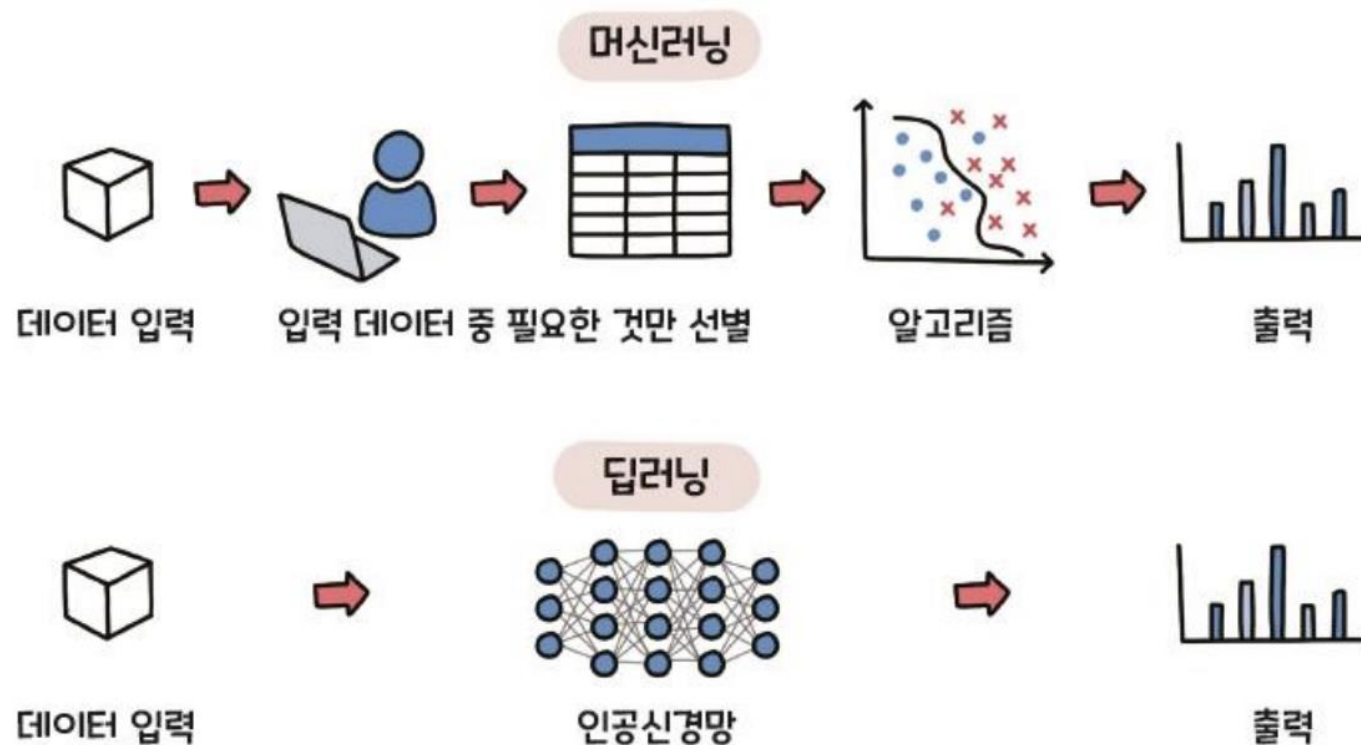


그림 1-28 머신러닝과 딥러닝의 학습 차이

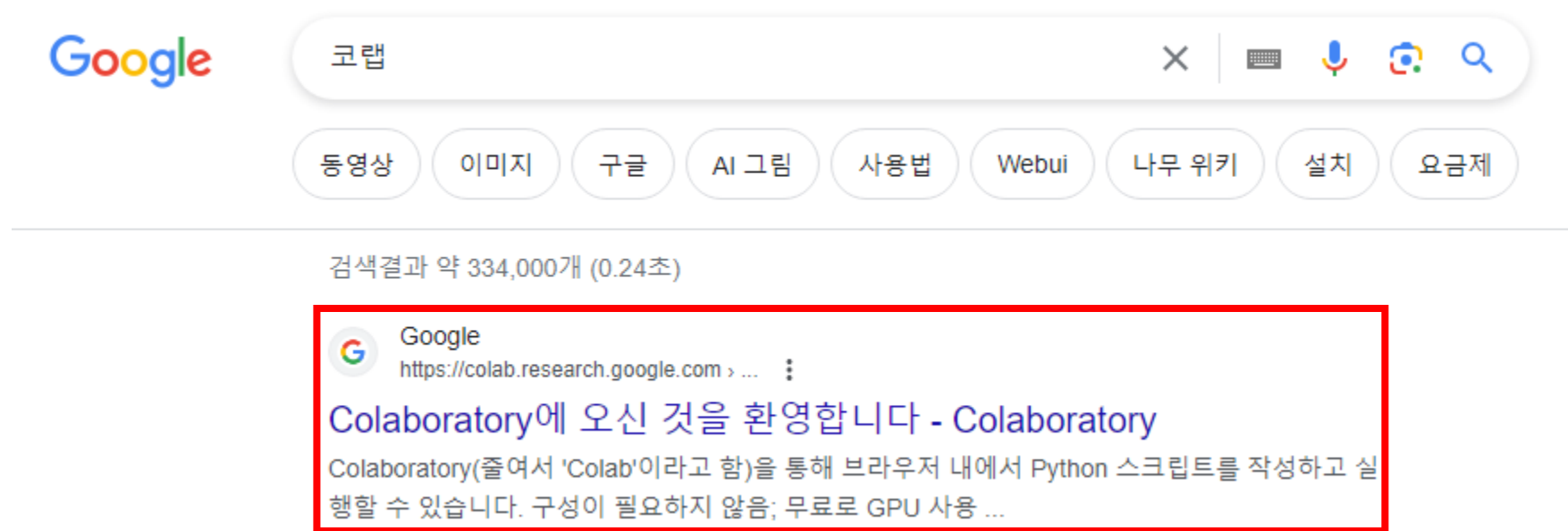
2. 환경 세팅

- **구글 계정** 준비
- 사용할 IDE : Colab(Google)



2. 환경 세팅

구글에 로그인 하기



2. 환경 세팅



Colaboratory에 오신 것을 환영합니다

파일 수정 보기 삽입 런타임 도구 도움말

새 노트
노트 열기 Ctrl+O
노트 업로드
이름 바꾸기
Drive에 사본 저장
GitHub Gist로 사본 저장
GitHub에 사본 저장
저장 Ctrl+S
업데이트 기록
다운로드
인쇄 Ctrl+P

+ 코드 + 텍스트 | Drive로 복사

Colab 시작 페이지

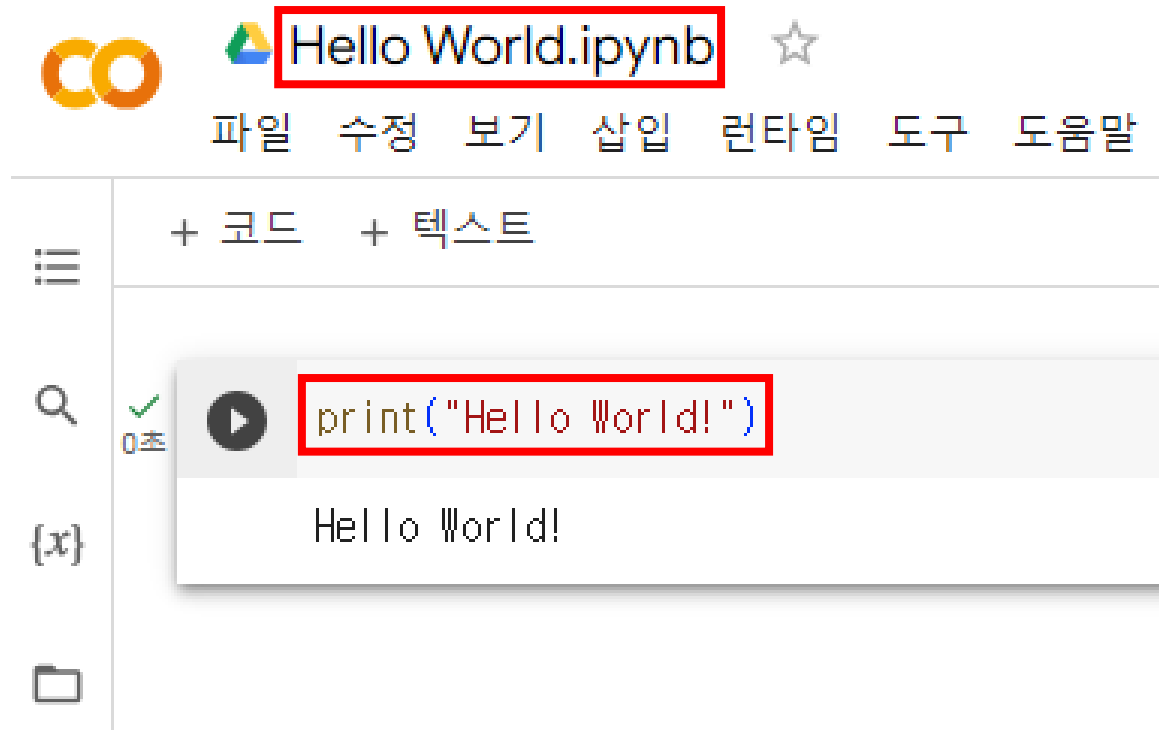
Colab에 이미 익숙하다면 이 동영상을 통해 양방향 테이블, 코드 실행 기록 보기, 명령어 팔레트에 관해 알아보세요.



Colab이란?


Colaboratory(줄여서 'Colab'이라고 함)을 통해 브라우저 내에서 Python 스크립트를 작성하고 실행할 수 있습니다.

2. 환경 세팅



The screenshot displays the Google Colab web interface. At the top left is the Colab logo. Next to it is the notebook title 'Hello World.ipynb', which is highlighted with a red rectangular box. To the right of the title is a star icon. Below the title is a horizontal menu with the following items: '파일' (File), '수정' (Edit), '보기' (View), '삽입' (Insert), '런타임' (Runtime), '도구' (Tools), and '도움말' (Help). Below the menu is a toolbar with '+ 코드' (Add Code) and '+ 텍스트' (Add Text) buttons. On the left side of the interface is a vertical sidebar containing icons for a menu, search, a variable '{x}', and a folder. The main workspace shows a single code cell. The code cell has a play button icon on the left, a green checkmark and '0초' (0 seconds) status indicator, and the code 'print("Hello World!")' which is highlighted with a red rectangular box. Below the code cell, the output 'Hello World!' is displayed.

2. 환경 세팅

 **드라이브**

+ 새로 만들기

내 드라이브

컴퓨터

공유 문서함

최근 문서함

중요 문서함

스팸

휴지통

저장용량

15GB 중 8.27GB 사용

추가 저장용량 구매

🔍 드라이브에서 검색

👤

내 드라이브 > Colab Notebooks ▾


유형 ▾


사람 ▾

수정 날짜 ▾

이름 ↑

소유자

 Hello World.ipynb

 나

Tip! 마크다운

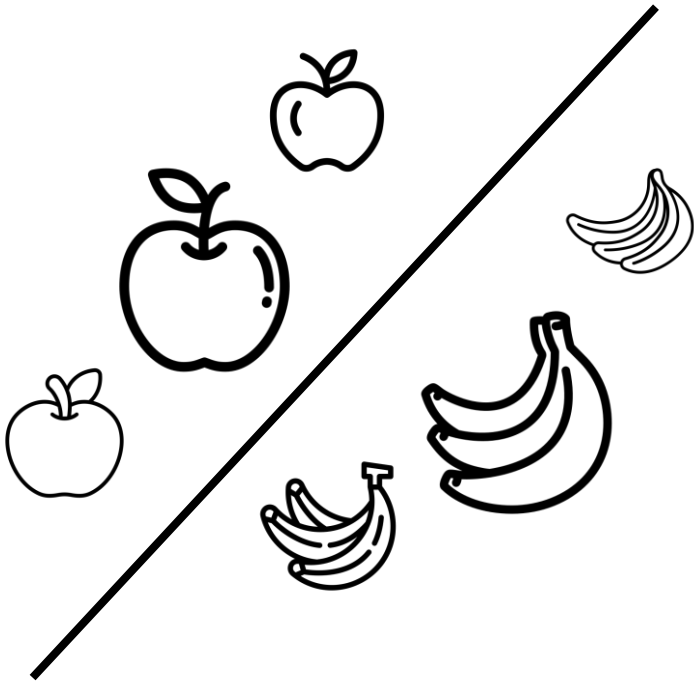
• 텍스트 셀에서 사용

텍스트 셀에 사용할 수 있는 마크다운

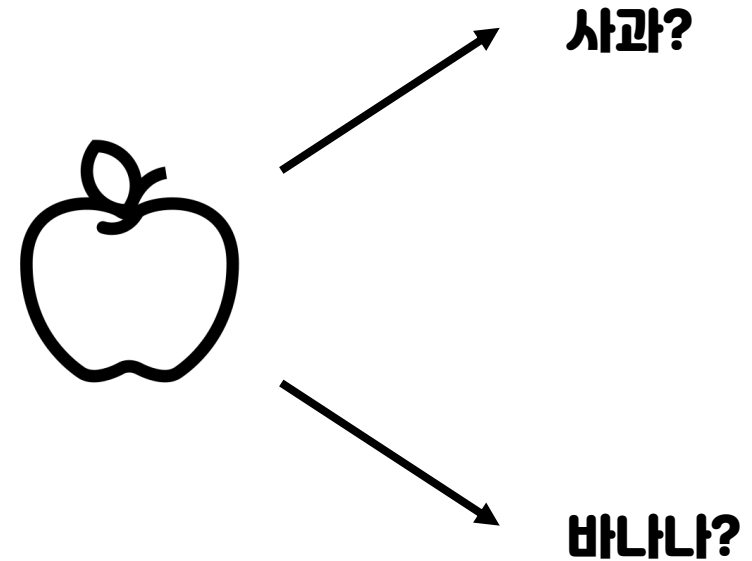
마크다운 형식	설명	예제
# 제목1	<h1> 태그와 동일합니다.	제목1
## 제목2	<h2> 태그와 동일합니다.	제목2
### 제목3	<h3> 태그와 동일합니다.	제목3
#### 제목4	<h4> 태그와 동일합니다.	제목4
##### 제목5	<h5> 태그와 동일합니다.	제목5
혼공머신	굵게 씁니다.	혼공신
혼공머신 _혼공머신_	가울임 꼴로 씁니다.	혼공신
~~혼공머신~~	취소선을 추가합니다.	혼공신
`print("Hello World!")`	백틱 기호를 사용해 코드 서체로 씁니다.	<code>print("Hello World!")</code>
> 혼공머신	들여쓰기합니다. 여러 단계를 들여쓸 수 있습니다.	혼공신
* 혼공머신 - 혼공머신	글머리 기호 목록을 만듭니다.	• 혼공신
[한빛미디어](http://www.hanbit.co.kr/)	링크를 만듭니다.	한빛미디어
![한빛미디어](http://www.hanbit.co.kr/images/common/logo_hanbit.png)	이미지를 추가합니다.	 한빛출판네트워크
\$ y = x \times z \$	레이텍을 추가합니다.	$y = x \times z$

note 레이텍(LaTeX)은 수식, 그래프, 다이어그램 등을 그리는 데 유용한 문서 저작도구로 보통 논문 작성에 많이 사용됩니다.

3. 머신러닝 기초



사과, 바나나는 **클래스(class)**



분류(classification)

특히 두 개의 클래스 중 하나를 고르는 것은 **이진 분류(binary classification)**

3. 머신러닝 기초



산점도
(scatter)



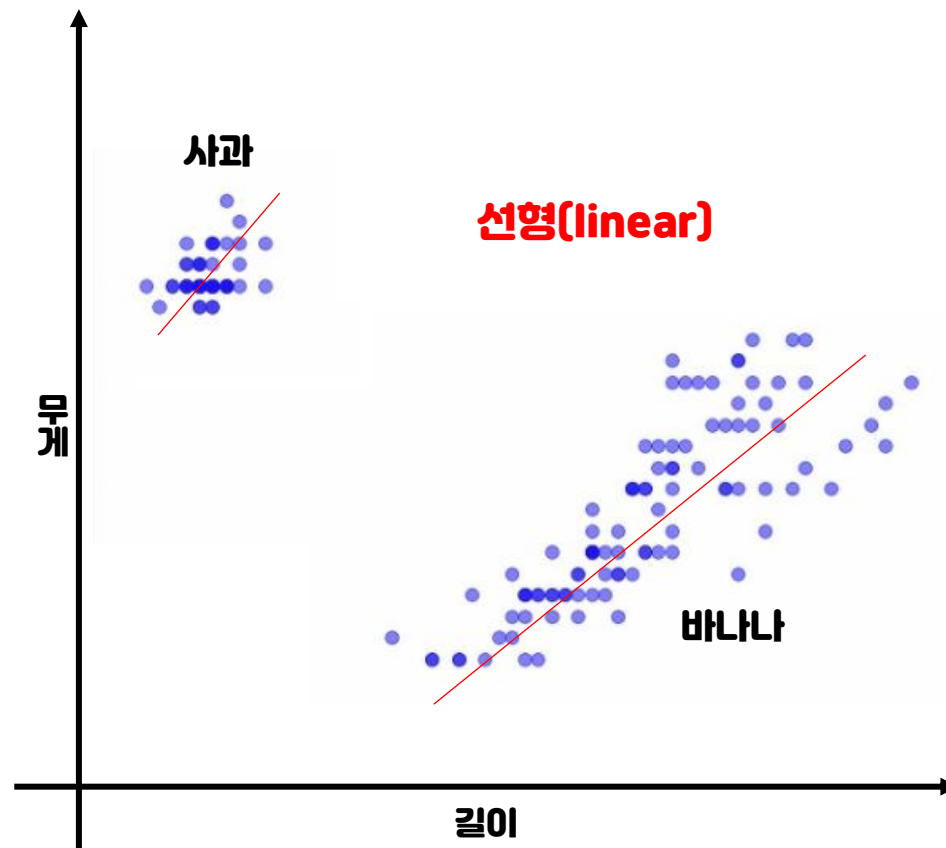
길이 : 10cm

20cm

무게 : 300g

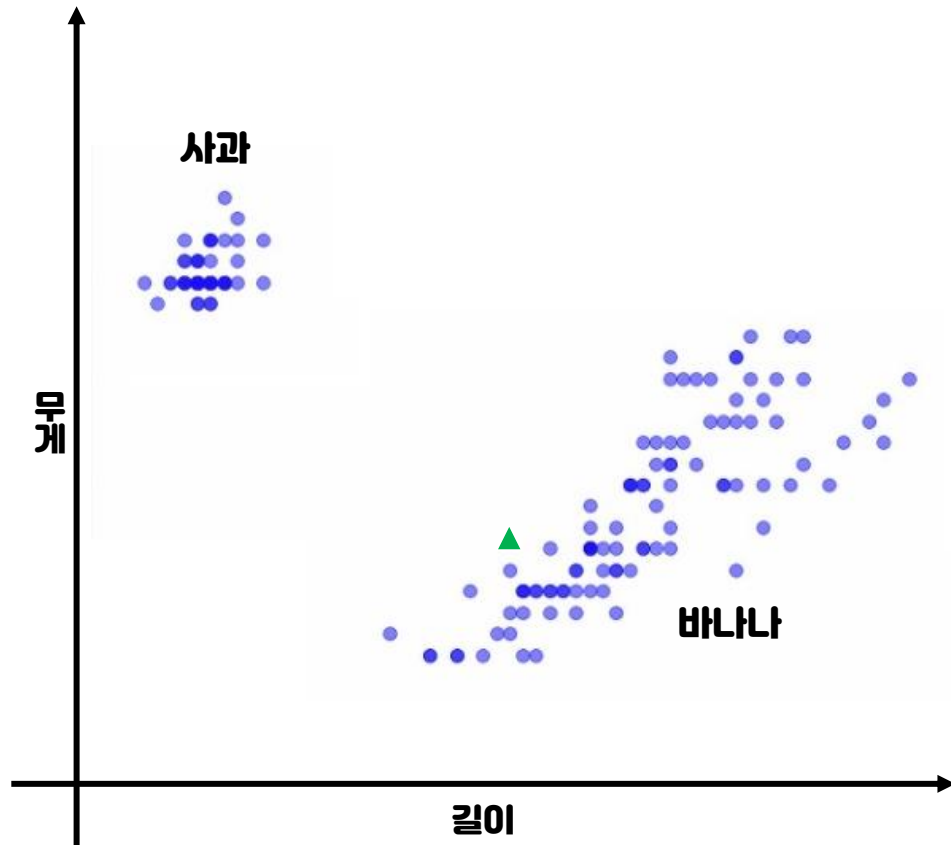
200g

길이와 무게 같은 특징은 **특성(feature)**





k-NN 알고리즘 (k-최근접이웃 알고리즘)



k-NN(k-Nearest Neighbor) 알고리즘

어떤 데이터에 대한 답을 구할 때,
주변의 **k개의 가장 가까운 데이터**를 보고
다수를 차지하는 것을 정답으로 판단

=> **초록색 삼각형**은? **바나나!**

k-NN 알고리즘 단점

- **k-NN 알고리즘 : 데이터를 모두 가지고 있는 것이 전부
=> 데이터가 아주 많은 경우 사용이 어려움**